Cell Reports
Article

Cell PRESS
Open ACCESS

Long-Timescale Dynamics and Regulation of Sec-Facilitated Protein Translocation

Bin Zhang[1] and Thomas F. Miller III[1,*]
[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA
*Correspondence: tfm@caltech.edu
http://dx.doi.org/10.1016/j.celrep.2012.08.039

SUMMARY

We present a coarse-grained modeling approach that spans the nanosecond- to minute-timescale dynamics of cotranslational protein translocation. The method enables direct simulation of both integral membrane protein topogenesis and transmembrane domain (TM) stop-transfer efficiency. Simulations reveal multiple kinetic pathways for protein integration, including a mechanism in which the nascent protein undergoes slow-timescale reorientation, or flipping, in the confined environment of the translocon channel. Competition among these pathways gives rise to the experimentally observed dependence of protein topology on ribosomal translation rate and protein length. We further demonstrate that sigmoidal dependence of stop-transfer efficiency on TM hydrophobicity arises from local equilibration of the TM across the translocon lateral gate, and it is predicted that slowing ribosomal translation yields decreased stop-transfer efficiency in long proteins. This work reveals the balance between equilibrium and nonequilibrium processes in protein targeting, and it provides insight into the molecular regulation of the Sec translocon.

INTRODUCTION

The Sec translocon is a central component of the cellular machinery for targeting and delivering nascent proteins. Ubiquitous across all kingdoms of life, it is a protein-conducting channel that facilitates recognition of integral membrane protein domains and the establishment of integral membrane protein topology. Extensive structural (Van den Berg et al., 2004; Egea and Stroud, 2011; Frauenfeld et al., 2011; Tsukazaki et al., 2008; Zimmer et al., 2008), biochemical (Bonardi et al., 2011; Cheng and Gilmore, 2006; Do et al., 1996; Duong and Wickner, 1998; Kim et al., 2002; Park and Rapoport, 2011), and genetic (Bieker and Silhavy, 1990; Junne et al., 2007; Smith et al., 2005) analysis has illuminated the role of the translocon in both cotranslational and posttranslational pathways.

Quantitative biological assays have revealed the sensitivity of the translocon to changes in molecular interactions and external driving forces. In particular, the groups of von Heine and White have identified a "translocon code" that relates the physicochemical properties of a nascent protein to its relative propensity for translocation versus membrane integration (Hessa et al., 2005, 2007); this substrate-determined regulation of stop-transfer efficiency is consistent with an apparent equilibrium partitioning of the nascent protein between hydrophobic and hydrophilic environments. Furthermore, Spiess and colleagues have demonstrated the dependence of membrane protein topogenesis on protein sequence, mutations in the translocon channel, and the rate at which the nascent protein is cotranslationally inserted into the channel (Goder and Spiess, 2003; Higy et al., 2005; Junne et al., 2007), which provides striking evidence for the role of kinetic effects in the regulation of Sec-facilitated membrane integration. However, no coherent approach currently exists to explore the mechanistic basis for these observed kinetic effects. Nor is it clear how to reconcile the apparent role of equilibrium partitioning in the work of von Heine and White with the effects of nonequilibrium (i.e., kinetic) regulation in the work of Spiess and colleagues. The current study aims to address these challenges and to establish fundamental connections between previously disparate experimental studies of Sec-facilitated protein translocation and integral membrane protein topogenesis.

The development of a unified, mechanistic understanding of Sec-facilitated protein targeting is hindered by the complex and important roles of collaborating molecular motors, large-scale conformational changes in the translocon, and the crowded molecular environment of the channel interior. Computer simulation studies provide a useful approach to understanding the translocon by connecting high-resolution structures to its detailed molecular interactions and dynamics (Bondar et al., 2010; Gumbart et al., 2011; Gumbart and Schulten, 2006, 2007; Haider et al., 2006; Tian and Andricioaei, 2006; Zhang and Miller, 2010, 2012). Yet the biological timescales for cotranslational protein translocation (i.e., minutes) vastly exceed the reach of atomistic MD simulations (Rychkova et al., 2010), and the large number of trajectories needed to explore the parameter space of protein sequence and translation rate with statistical significance ($\sim 10^5$ in the current study) dramatically constrains the computational cost of applicable simulation methods. New approaches are needed to bridge the hierarchy of timescales in Sec-facilitated protein translocation and membrane integration and to identify the mechanisms that govern these fundamental cellular processes.

In this study, we develop a coarse-grained (CG) model that enables simulation of the translocon and its associated macromolecular components on timescales beyond the scope of

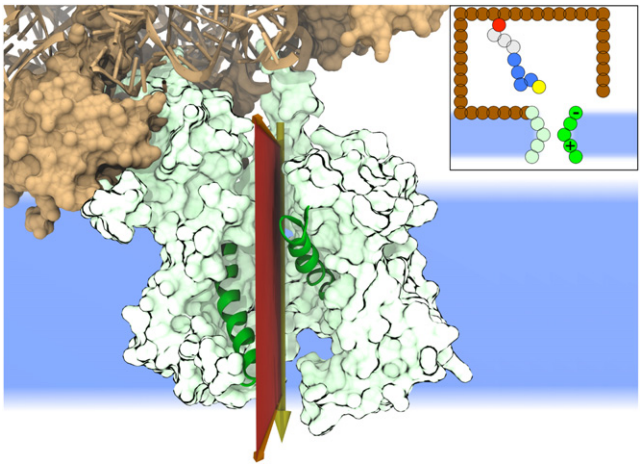

Figure 1. Structural Features of the Cotranslational Sec Machinery
The ribosome (brown) is shown in complex with the Sec translocon (green). The CG model projects the protein nascent chain dynamics onto the plane (red) that intersects the translocon channel axis and that bisects the lateral gate (LG) helices (dark green). (inset) The CG model includes beads for the translocon (green), the ribosome (brown), and the protein nascent chain. The LG helices are shown in dark green, the ribosome exit channel is shown in red, and the lipid membrane is shown in blue. The nascent chain is composed of beads for the SP (yellow and blue) and the mature domain (gray).

previously employed methodologies. The model explicitly describes the configurational dynamics of the nascent protein chain, conformational gating in the Sec translocon, and the slow dynamics of ribosomal translation (Figure 1). We use the model to perform minute-timescale CG trajectories to investigate the role of the Sec translocon in governing both stop-transfer efficiency (i.e., propensity of TM to undergo integration into the cell membrane versus secretion across the membrane) and integral membrane protein topogenesis (i.e., the propensity of TM to undergo membrane integration in the N_{cyt}/C_{exo} orientation versus the N_{exo}/C_{cyt} orientation). These simulations provide a direct probe of the mechanisms, kinetics, and regulation of Sec-facilitated protein translocation and membrane integration. Analysis of the full ensemble of nonequilibrium CG trajectories reveals the molecular basis for experimentally observed trends in integral membrane protein topogenesis and TM stop-transfer efficiency; it demonstrates the role of competing kinetic pathways and slow conformational dynamics in Sec-facilitated protein targeting; and it provides experimentally testable predictions regarding the long-timescale dynamics of the Sec translocon.

RESULTS

Signal Orientation and Protein Topogenesis

Signal peptide (SP) orientation is a determining factor in integral membrane protein topogenesis (Goder and Spiess, 2001). The orientation of N-terminal signals help to establish the topology of multidomain integral membrane proteins and to dictate whether N-terminal or C-terminal domains undergo translocation across the membrane. Biochemical studies have established the dependence of SP orientation upon a range of factors, including SP flanking charges (Beltzer et al., 1991; Parks and Lamb, 1991), SP hydrophobicity (Harley et al., 1998; Hikita and Mizushima, 1992; Wahlberg and Spiess, 1997), protein mature domain length (MDL) (Goder and Spiess, 2003), and the ribosomal translation rate (Goder and Spiess, 2003). In this section, we employ the CG model to directly simulate cotranslational protein integration and to determine the molecular mechanisms that give rise to these experimentally observed relations.

Direct Simulation of Cotranslational Protein Integration

We consider the process in which cotranslational integration of a signal anchor protein yields either the type II (N_{cyt}/C_{exo}) or type III (N_{exo}/C_{cyt}) orientation of the uncleaved SP domain; this nomenclature for the orientation of single-spanning membrane proteins follows earlier work (Goder and Spiess, 2001). Figure 2 illustrates the simulation protocol, with the N-terminal SP domain shown in blue and yellow.

Following previous experimental work (Goder and Spiess, 2003), we consider the integration of proteins that vary with respect to both SP sequence and MDL. The SP is composed of either a canonical sequence of CG beads (RL_4E), a sequence in which the positive charge on the N-terminal group is eliminated (QL_4E), or a sequence with enhanced SP hydrophobicity (RL_6E). To model the hydropathy profile of the engineered protein H1ΔLeu22 studied by Goder and Spiess (2003) (Figure S1), we consider proteins that include a hydrophilic mature domain with a hydrophobic patch near the SP; specifically, we model the protein mature domain using the Q_5LQ_n sequence of CG beads, such that the total peptide length ranges from 30–80 beads (90–240 residues [res]). The sensitivity of protein topology to hydrophobic patches on the mature domain is examined in Figure S2A.

CG trajectories are continued until the protein nascent chain reaches either type II or type III integration. Depending upon the rate of ribosomal translation and the MDL, each CG trajectory thus ranges from 2–20 s of simulation time; the corresponding CPU time required to perform each trajectory is approximately 0.2–10 hr. Each data point in Figures 3A–3C is obtained by averaging the results of at least 600 independent CG trajectories. Full details of the simulation protocol are provided in Extended Experimental Procedures. Representative trajectories are illustrated in Movies S1 and S2.

Figures 3A–3C present the fraction of peptides that are calculated to undergo type II integration as a function of protein MDL. In each case, the CG model predicts a strong dependence of SP topology on the length of the protein mature domain, with a fast rise in the type II integration fraction at short lengths plateauing to a fixed value at longer MDL. The CG model also finds significant dependence of signal topology on the SP charge distribution (Figure 3A), SP hydrophobicity (Figure 3B), and ribosomal translation rate (Figure 3C). Each of these trends is in striking agreement with the findings of Goder and Spiess (2003); in addition to the crossover from strong to weak dependence of the signal topology with increasing MDL, the experimental study likewise reports type II integration to be reduced with the removal of positively charged N-terminal groups, more hydrophobic SP sequences, and faster protein insertion (see

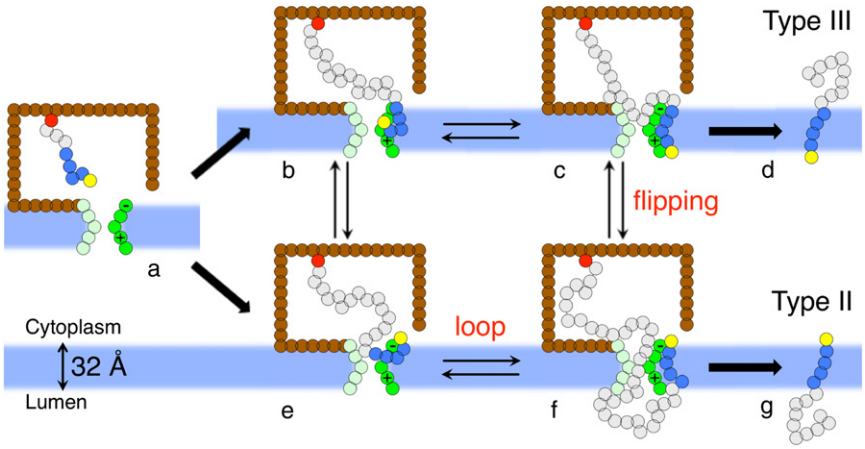

Figure 2. Kinetic Pathways for Type II and Type III Membrane Integration of Signal Anchor Proteins Obtained from Direct CG Simulations

The coloring scheme is described in Figure 1. States *a*–*g* observed in the mechanism are described in the text.

also Figures S3A and S3B). Figures S2C–S2F provide additional tests and comparisons of the CG model against protein topogenesis experiments, analyzing factors that include negative N-terminal charges, elongated N-terminal domains, charge mutations on the translocon, and charged patches on the nascent-protein mature domain. In the following, we use the CG simulations to enable the detailed analysis of the insertion dynamics and to determine the mechanistic origin of these various trends.

Competition between Kinetic Pathways Governs Topogenesis

Inspection of the ensemble of CG trajectories reveals multiple kinetic pathways by which the protein nascent chain achieves type II or type III integration (Figure 2). During early-stage protein insertion, the SP typically binds at the lateral gate (LG) in one of two conformations, either with its N terminus buried inside the translocon (state *b*) or exposed to the membrane (state *e*); similar conformations have been observed in microsecond-timescale, all-atom MD simulations of early-stage peptide insertion (Figure S3C) (Zhang and Miller, 2012). From state *e*, further insertion of the nascent chain yields state *f*, in which the SP assumes the N_{cyt}/C_{exo} orientation; continued translocation of the mature domain in this orientation eventually leads to type II integration. From state *b*, further insertion leads to state *c*, in which the SP assumes the N_{cyt}/C_{exo} orientation; this orientation does not directly facilitate mature domain translocation, without which the protein assumes type III integration. Slow transitions between states *c* and *f* are also observed in many trajectories; this conformational change, in which the SP "flips" between type III and type II integration topologies, is found to lie at the heart of many of the trends in Figures 3A–3C.

To analyze the flow of trajectories among these competing mechanisms, the CG trajectories are categorized according to the chronology with which they pass through the states *a*-*g* in Figure 2. Each trajectory is associated with either type III mechanism (*a*-*b*-*c*-*d*), the type II loop mechanism (*a*-*e*-*f*-*g*), or the type II flipping mechanism (*a*-*b*-*c*-*f*-*g*). We emphasize that trajectories need not pass irreversibly through these states. Trajectories that visit state *c* prior to type II integration are associated with the flipping mechanism, whereas any other trajectory that reaches type II integration is associated with the loop mechanism; all remaining trajectories are associated with the type III mechanism. The definition for state *c* in terms of the coordinates of the model is presented in Extended Experimental Procedures. Figure 3D presents the fraction of trajectories passing through each of these competing mechanisms, and it compares the effect of SP sequence and translation rate on the mechanism of integration. A total protein nascent chain length of 210 residues is considered for all cases in this figure.

Differences between the RL_4E and QL_4E data sets in Figure 3D help to explain the shift between the two corresponding data sets in Figure 3A. For the canonical SP sequence (RL_4E), Figure 3D shows that CG trajectories predominantly follow the type II loop mechanism for integration. However, upon mutating the SP sequence with respect to the number of charged residues (QL_4E), the type II flipping mechanism and the type III mechanism become more prevalent. Removal of the N-terminal charge group diminishes the electrostatic stabilization of the SP in the N_{cyt}/C_{exo} orientation. The CG trajectories are thus less likely to visit states *e* and *f*, which are on pathway for type II loop integration, in favor of states *b* and *c*, which are on pathway for both type II flipping and type III integration. Interestingly, the flipping mechanism allows for significant compensation of the type II integration fraction upon mutation of the charge group; the effect of the SP sequence mutation on the flow of CG trajectories (Figure 3D) is thus much greater than the corresponding effect on the final branching ratio between type II and type III integration (Figure 3A). The simulations reveal a competition between electrostatic stabilization and SP reorientation kinetics that contributes to the well-known "positive-inside rule" for integral membrane protein topology (Goder and Spiess, 2003; von Heijne, 1986). Furthermore, these results suggest that hindering the $c \rightarrow f$ flipping transition, perhaps via small molecule binding (Garrison et al., 2005; Maifeld et al., 2011), may lead to a larger effect on the type II integration fraction than is observed with N-terminal charge mutation.

Comparison of the data for the RL_4E and RL_6E sequences in Figure 3D explains the shift between the two corresponding data sets in Figure 3B. Figure 3D shows that increasing the hydrophobicity of the SP reduces the flow of integration trajectories through the type II loop mechanism. As before, this can be attributed to changes in the stability of states along the competing kinetic pathways. Increasing the hydrophobicity of the SP sequence significantly stabilizes SP configurations in state *b*, which favorably expose the hydrophobic segment to the membrane, instead of configurations in state *e*, which bury

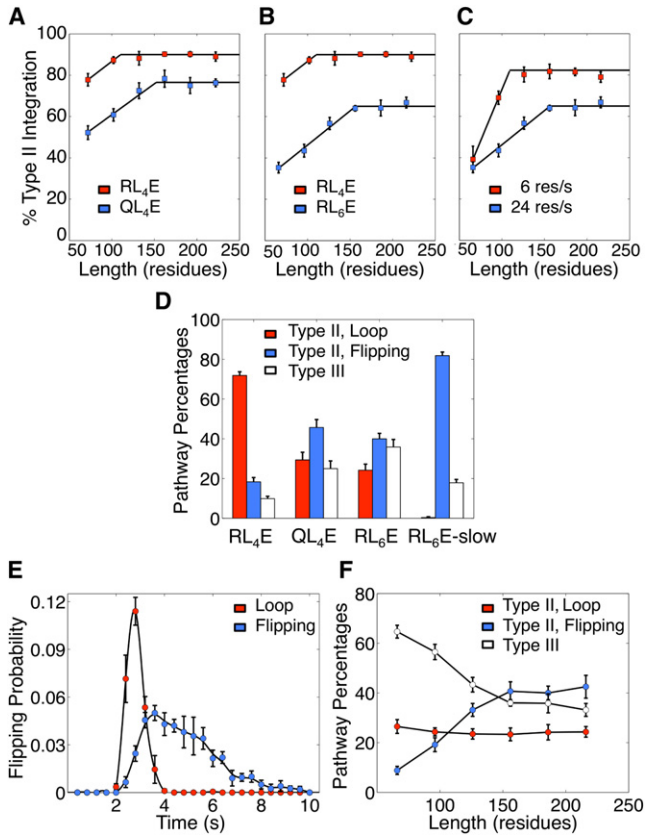

Figure 3. CG Simulation Results for Integral Membrane Protein Topogenesis

(A–C) Fraction of type II integration as a function of protein MDL, with data sets that vary with respect to (A) SP charge distribution, (B) SP hydrophobicity, and (C) ribosomal translation rate.

(D) Fraction of CG trajectories that follow the type II loop pathway (red), type II flipping pathway (blue), and the type III pathway for membrane integration (white).

(E) The distribution of arrival times for CG trajectories at state *f* of type II integration via the loop pathway (red) and the flipping pathway (blue).

(F) MDL dependence of the fraction of CG trajectories that follow each integration mechanism. Unless otherwise specified, error bars throughout the paper represent the SD of the mean.

See also Figures S1, S2, and S3.

the hydrophobic segment inside the translocon. This effect draws trajectories away from the loop mechanism (Figure 3D) and leads to decreased type II integration (Figure 3B).

Differences between the RL_6E and RL_6E-slow data sets in Figure 3D help to explain the shift between the two corresponding data sets in Figure 3C. Slowing the rate of ribosomal translation in proteins from 24 res/s to 6 res/s causes the CG trajectories to shift almost entirely to a type II flipping mechanism. These differences are remarkable since they involve no change in the interactions of the system; the shifts in SP topology (Figure 3C) and integration mechanism (Figure 3D) with protein translation rate are purely kinetic effects. With slower translation, partially translated protein nascent chains have more time to undergo conformational sampling and are more likely to visit state *c*; it is therefore expected that Figure 3D shows type II loop integration decreases in favor of combined type II flipping integration and type III integration. However, the corresponding decrease in type III integration is more surprising.

The decrease in type III integration upon slowing translation arises from the important role of the flipping transition from state *c* to state *f*, which enables the nascent chain to reach the more thermodynamically favorable configurations associated with the N_{cyt}/C_{exo} SP orientation. Figure 3E plots the distribution of arrival times at state *f* for trajectories that follow either the type II loop mechanism (red) or the type II flipping mechanism (blue). Trajectories complete the loop mechanism relatively quickly, whereas the timescale for flipping persists as long as 10 s. The flipping transition thus introduces a slow timescale for conformational dynamics that couples to the dynamics of ribosomal translation. Slowing ribosomal translation provides more time for the nascent chain to undergo flipping; this purely kinetic effect enhances type II integration in Figure 3C.

The final trend left to explain in Figures 3A–3C is the dependence of the type II integration fraction on the MDL. For every data set, the type II integration fraction increases with MDL before plateauing to a constant value. Figure 3F elucidates this trend by presenting how the insertion mechanism varies with MDL; the percentage of CG trajectories following each mechanism is calculated as in Figure 3D.

With increasing MDL (Figure 3F), the fraction of trajectories following the type II loop mechanism remains relatively unchanged, whereas the prevalence of type II flipping increases at the expense of the type III mechanism. As was seen from Figure 3E, trajectories commit to the type II loop mechanism relatively early during insertion, prior to the full completion of ribosomal translation; it follows that increasing the MDL will have little effect on the fraction of trajectories following this mechanism. Furthermore, the tradeoff in Figure 3F between the type II flipping and type III mechanisms occurs for the same reason as was discussed for slowed ribosomal translation; increasing the MDL in Figure 3F provides more time for the tethered nascent chain to undergo the slow flipping transition from state *c* to the thermodynamically favored state *f*. At long MDL, the crowded environment in the ribosome-translocon junction causes nascent chain configurations in state *c* to be driven into state *d* before they can undergo the flipping transition; this causes the fraction of type II flipping trajectories to cease rising in Figure 3F, such that the relative fraction of type II flipping and type III trajectories approaches a constant value. The results in Figure 3F correspond to the particular case of the RL_6E SP sequence and the 24 res/s translation rate; however, the trends are general and explain the MDL dependence of the type II integration fraction in Figures 3A–3C.

Loop versus Flipping Mechanisms

Observation of competing pathways for type II integration is an unexpected and significant feature of the CG simulations presented here. Both the loop and flipping mechanisms for SP integration have been proposed in previous experimental studies (Devaraneni et al., 2011; Goder and Spiess, 2003; Rapoport et al., 2004; Shaw et al., 1988), although the possible role of peptide sequence and ribosomal translation rate in converting

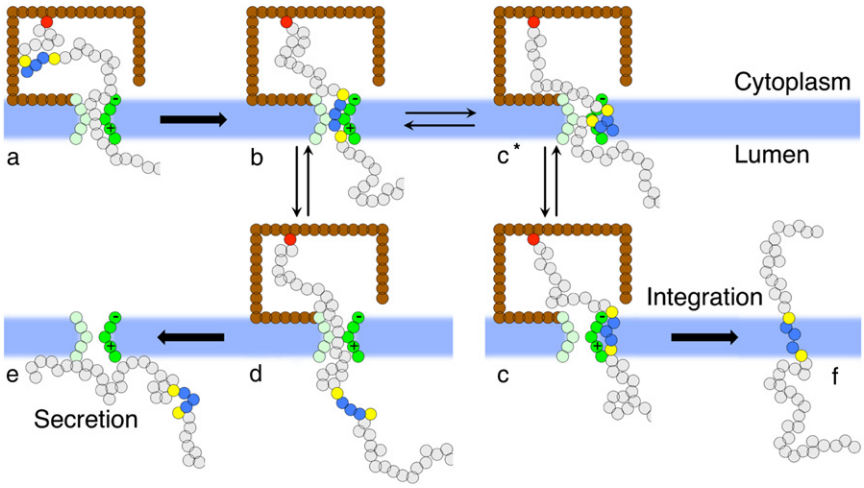

Figure 4. Kinetic Pathways for Cotranslational Protein Translocation and Membrane Integration Obtained from Direct CG Simulations

The H-domain of the protein nascent chain is shown in blue and yellow. The full N-terminal anchor domain of the protein nascent chain is not shown here; the full system is shown in Figure S4. States *a–f* observed in the mechanism are described in the text.

between these mechanisms has not been emphasized. Experimental support for the loop mechanism includes evidence that the protein nascent chain remains enclosed within the ribosome-translocon junction during the establishment of SP orientation (Beckmann et al., 2001). Indeed, nascent proteins are found to be protected from cytosolic fluorescent quenching agents (Crowley et al., 1994, 1993) or proteases (Jungnickel and Rapoport, 1995; Rutkowski et al., 2001) in some systems, although proteins with more hydrophobic SP sequences are found to exhibit protease degradation in translation-stalled intermediates (Rutkowski et al., 2001). The loop mechanism is also consistent with observations that type II integration is uninhibited by inclusion of bulky N-terminal domains in the protein nascent chain sequence (Denzer et al., 1995; Shaw et al., 1988). On the other hand, Spiess, Rapoport, and colleagues have proposed the flipping mechanism for type II integration to explain observed trends in SP topogenesis (Goder and Spiess, 2003; Higy et al., 2005; Rapoport et al., 2004); and direct evidence in support of the flipping mechanism has recently been reported (Devaraneni et al., 2011) under the assumption that translation-stalled intermediates of the ribosome/translocon/nascent-chain complex reflect the kinetic pathway for membrane integration. The observed coexistence of the loop and flipping mechanisms in our CG simulations helps to reconcile these experimental findings, and it provides a basis for understanding the competing influences of SP hydrophobicity, SP charge distribution, MDL, and ribosomal translation rate in regulating Sec-facilitated type II and type III protein integration.

In assessing the role of the type II flipping mechanism in physiological systems, we note that many naturally occurring proteins exhibit longer N-terminal domains and less hydrophobic SP than the protein sequences considered in both here and in the work of Goder and Spiess (2003). As discussed previously, Figure 3D reveals that decreasing SP hydrophobicity leads to a decrease in the fraction of trajectories undergoing the type II flipping mechanism. Furthermore, CG simulations performed using protein nascent chain sequences with longer N-terminal domains (Figure S2C) reveal a corresponding decrease in the fraction of trajectories that exhibit the type II flipping mechanism.

Regulation of Stop-Transfer Efficiency

In addition to facilitating the translocation of proteins across the phospholipid membrane, the Sec translocon plays a key role in determining whether nascent protein chains become laterally integrated into the membrane (Rapoport et al., 2004). Strong correlations between the hydrophobicity of a TM and its stop-transfer efficiency have led to the suggestion of an effective two-state partitioning of the TM between the membrane interior and a more aqueous region (Heinrich et al., 2000; Hessa et al., 2005). However, models for this process based purely on the thermodynamic partitioning of the TM do not account for the experimentally observed dependence of stop-transfer efficiency on the length of the protein nascent chain (Hessa et al., 2003), nor would such models anticipate any change in TM partitioning upon slowing ribosomal translation. Furthermore, recent theoretical (Zhang and Miller, 2010) and experimental work (Junne et al., 2010) point out that the observed correlations between stop-transfer efficiency and substrate hydrophobicity can also be explained in terms of a kinetic competition between the secretion and integration pathways under the substrate-controlled conformational gating of the translocon. To further elucidate the mechanism of Sec-facilitated regulation of protein translocation and membrane integration, we employ the CG model to directly simulate cotranslational stop-transfer regulation and to analyze the role of competing kinetic and energetic effects.

Direct Simulation of Cotranslational TM Partitioning

Following recent experimental studies (Hessa et al., 2005, 2007; Junne et al., 2010), we consider the cotranslational partitioning of a stop-transfer TM (i.e., the H-domain) where the protein nascent chain topology is established by an N-terminal anchor domain. Stop-transfer efficiency is defined as the fraction of translated proteins that undergo H-domain membrane integration, rather than translocation. Figure 4 illustrates the simulation protocol, with the H-domain shown in blue; see also Figure S4, which shows the full system including the anchor domain.

The translated protein sequence is comprised of three components, including the N-terminal anchor domain, the H-domain, and the C-terminal tail domain. In all simulations, the N-terminal anchor domain includes 44 type-Q CG beads that link the H-domain to an anchor TM that is fixed in the N_{cyt}/C_{exo} orientation (Figure S4). The H-domain is comprised of the sequence

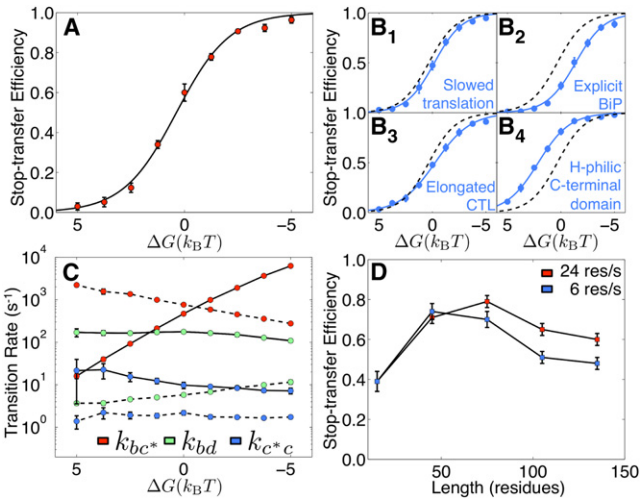

Figure 5. CG Simulation Results for TM Partitioning

(A) Stop-transfer efficiency as a function of H-domain hydrophobicity.

(B) Dependence of stop-transfer efficiency upon (B_1) slowing ribosomal translation rate from 24 res/s to 6 res/s, (B_2) including explicit lumenal BiP binding, (B_3) increasing the CTL from 75 residues to 105 residues, and (B_4) replacing the hydrophobic beads in the protein C-terminal domain with hydrophilic beads; in each subpanel, the dashed line corresponds to the sigmoidal fit of the data in (A).

(C) Equilibrium transition rates between the states in Figure 4 as a function of H-domain hydrophobicity. For each color, the forward rate is indicated with the solid line, and the reverse rate is indicated with dashed line.

(D) Dependence of stop-transfer efficiency on CTL and the ribosomal translation rate, obtained for protein sequences with H-domain transfer FE of $\Delta G = -1.25 k_B T$.

Error bars represent the SD of the mean. See also Figures S1 and S5.

PX_3P, where the X-type CG beads have variable hydrophobicity. The C-terminal domain includes a hydrophilic sequence of CG beads with periodic hydrophobic patches (poly-Q_5V), following the hydrophobicity profile of the dipeptidyl aminopeptidase B (DPAPB) protein studied by Junne and colleagues (Figure S1) (Junne et al., 2010).

Stop-transfer efficiency is studied as a function of the hydrophobicity of the H-domain, the C-terminal tail length (CTL), and the ribosomal translation rate. We consider CTL in the range of 5–45 beads (15–135 residues), and we consider water-membrane transfer free energies for the H-domain in the range of $\Delta G/k_B T = [-5, 5]$, where ΔG corresponds to the sum over the individual transfer free energies of the CG beads in the H-domain.

CG trajectories are initialized with the H-domain occupying the ribosome-translocon junction, prior to translation of the C-terminal domain (Figure 4, state *a*). Each CG trajectory is terminated after full translation of the protein C-terminal domain, either when the H-domain integrates into the membrane and diffuses a distance of 16 nm from the translocon or when both the H-domain and the C-terminal domain fully translocate into the lumenal region. The N-terminal anchor TM of the protein nascent chain is fixed at a distance of 20 nm from the translocon (Figure S4); the simulations thus assume that the H-domain membrane integration mechanism does not involve direct helix-helix contacts with the N-terminal anchor TM (Meindl-Beinker et al., 2006). Full details of the simulation protocol are provided in Extended Experimental Procedures. Representative trajectories are illustrated in Movies S3, S4, and S5.

Figure 5 presents the calculated dependence of stop-transfer efficiency on the hydrophobicity of the H-domain, the length and hydrophobicity of the protein C-terminal domain, and the ribosomal translation rate. Each data point in Figures 5A, 5B, and 5D is obtained from over 600 independent nonequilibrium CG trajectories; the simulation times for these trajectories span the range of 3–100 s. Figures S5A–S5C provide additional tests and comparisons of the CG model against stop-transfer experiments, analyzing factors that include charged residues flanking the H-domain, hydrophobic patches on the C-terminal domain, and changes in protein translocation time.

In Figure 5A, the stop-transfer efficiency is plotted as a function of the H-domain transfer FE, ΔG, for proteins with a CTL of 75 residues. The CG model recovers the experimentally observed (Hessa et al., 2005) sigmoidal dependence of stop-transfer efficiency on H-domain hydrophobicity. The black curve in the figure corresponds to the state population for a system in apparent two-state thermal equilibrium,

$$P_l(\Delta G) = (1 + \exp[-\beta\alpha\Delta G + \gamma])^{-1}, \qquad \text{(Equation 1)}$$

where $\alpha = -0.80$, $\gamma = 0.29$, and $\beta = (k_B T)^{-1}$ is the reciprocal temperature; see also Figure S5D. The physical origin of this sigmoidal dependence of the stop-transfer efficiency, as well as the physical interpretation of the parameters α and γ, is a focus of the following analysis.

Figure 5B presents the calculated relationship between stop-transfer efficiency and H-domain hydrophobicity in systems for which either the ribosomal translation rate is slowed from 24 res/s to 6 res/s (B_1), backsliding of the protein nascent chain is inhibited to explicitly model the effect of the lumenal BiP binding (B_2), the CTL is increased from 75 residues to 105 residues (B_3), or the hydrophobic patches (V-type beads) in the C-terminal domain are replaced with hydrophilic, Q-type beads (B_4). In each case, the integration probability preserves the sigmoidal dependence on ΔG, and the best-fit value for the parameter α in each case is remarkably unchanged from the case in Figure 5A. For the four cases presented in Figure 5B, fitting the simulation data to Equation 1 yields $\alpha = \{-0.77 \pm 0.08, -0.74 \pm 0.09, -0.60 \pm 0.06, -0.68 \pm 0.05\}$ and $\gamma = \{0.14 \pm 0.11, 1.0 \pm 0.19, -0.15 \pm 0.09, -1.44 \pm 0.13\}$; in each case, the 95% certainty threshold for the sigmoidal fit is also indicated (Sokal and Rohlf, 1994). Cases B_1–B_3 each lead to a decrease in the stop-transfer efficiency for a given value of ΔG (i.e., a rightward shift of the sigmoidal curve with respect to that obtained in Figure 5A), whereas decreasing the hydrophobicity of the C-terminal domain residues in case B_4 leads to an increase in stop-transfer efficiency.

The Origin of Hydrophobicity Dependence in TM Partitioning

Figure 4 introduces the primary mechanisms that the ensemble of CG trajectories are observed to follow in the simulations.

Along the pathway to membrane integration, trajectories pass through configurations for which the H-domain occupies the translocon channel (Figure 4, state *b*), the membrane-channel interface across the open LG (state *c**), and the membrane region outside of the translocon with the LG closed (state *c*); upon completion of translation and release of the protein nascent chain, it diffuses into the membrane to reach the integration product (state *f*). Along the pathway to protein translocation, trajectories also pass through state *b*, before proceeding to configurations in which the H-domain occupies the lumen with the C-terminal domain threaded through the channel (state *d*); upon completion of translation, the C-terminal domain is secreted through the channel, yielding the translocation product (state *e*). In addition to the dominant pathways depicted in Figure 4, minor pathways for translocation and integration are observed for very short and very long CTL (Figure S6A). Complete definitions for the states in Figure 4 in terms of the coordinates of the CG model are provided in Figure S7. We emphasize that trajectories do not irreversibly pass through the intermediate states in Figure 4; many trajectories backtrack repeatedly, starting down one pathway before finally proceeding down the other.

Figure 5C presents the equilibrium transition rates among the states in Figure 4, which are obtained from the frequency of interstate transitions in long CG trajectories of a protein nascent chain with a 75 residue C-terminal domain tethered at its C terminus to the ribosome exit channel. The calculation is repeated for proteins with a range of values for the H-domain hydrophobicity, ΔG. It is clear from the figure that partitioning of the H-domain across the LG of the translocon (i.e., forward and reverse transitions between states *b* and *c**) occurs on a faster timescale than most other transitions in the system. Furthermore, the rates k_{bc^*} and k_{c^*b} are strongly dependent on the hydrophobicity of the H-domain, whereas the other transition rates are only weakly dependent on ΔG.

The results in Figure 5C (as well as the more extensive kinetic analysis of the CG trajectories in Analytical Model for TM Partitioning) reveal the mechanistic origin of the observed sigmoidal dependence of TM partitioning on H-domain hydrophobicity (Figures 5A and 5B). The nascent protein H-domain achieves rapid, local equilibration (or partitioning) across the translocon LG; this partitioning is highly sensitive to the hydrophobicity of the H-domain, which gives rise to the characteristic sigmoidal dependence of the curves in Figures 5A and 5B and determines the value of the parameter α that appears in Equation 1. Moreover, rapid partitioning of the H-domain is kinetically uncoupled from slower steps in the mechanisms of integration and translocation, which leads to the insensitivity of α in fitting the various sets of data in Figures 5A and 5B. Kinetic and CTL effects in TM partitioning arise from competition among slower timescale processes in the secretion and integration pathways; these effects are manifest in parameter γ (Equation 1) and lead to lateral shifts of the sigmoidal curves in Figure 5B. We note that a mechanism involving local equilibration of the H-domain between the translocon and membrane interiors is consistent with the interpretation of recent experimental studies of stop-transfer efficiency (Junne et al., 2010; Ojemalm et al., 2011); however, the analysis presented here additionally reconciles the roles of both kinetic and thermodynamic effects in governing stop-transfer efficiency, and it provides a basis for understanding the lateral shifting of the sigmoidal curves both in Figure 5B and in possible future experiments.

Kinetic and CTL Effects in TM Partitioning

The direction of the lateral shifts of the curves in Figure 5B can also be understood from analysis of the CG trajectories. In part B_1, slowing the translation rate allows for better equilibration among the states *d* and *c* prior to release of the protein from the ribosome, leading to increased population of the thermodynamically favored state *d* and enhancement of the secretion product; Figure S6B demonstrates the relative increase of the nonequilibrium population in state *d* upon slowed ribosomal translation. In part B_2, the BiP motor enhances the secretion product by biasing against trajectories that backslide from state *d*. Part B_3 exhibits a combination of these two effects, with the elongated C-terminal domain allowing more time for the protein conformation to interconvert between states *d* and *c* prior to release from the ribosome (Figure S6B) and with a decreased rate of backsliding from state *d* with longer CTL (Figure S6C). Finally, part B_4 reveals that decreased hydrophobicity of the C-terminal domain residues leads to increased stop-transfer efficiency. Without hydrophobic patches, the C-terminal domain residues in the translocon channel do little to stabilize opening of the LG; therefore, once the system reaches state *c* along the pathway to membrane integration, it is less likely that the H-domain will return to the channel interior and then undergo secretion (Figure S5A).

Figure 5D provides a more complete view of the connection between CTL, ribosomal translation rate, and stop-transfer efficiency. At relatively long CTL ($\geq$75 res), stop-transfer efficiency decreases for longer proteins and for slower ribosomal translation, as was previously discussed in connection with Figures $5B_1$ and $5B_3$. However, at short CTL ($\leq$50 res), stop-transfer efficiency increases for longer proteins and exhibits no dependence on the ribosomal translation rate. In the short-CTL regimen, slowing ribosomal translation affords little additional time for the protein conformation to interconvert between states *d* and *c* prior to release from the ribosome (Figure S6B); there is thus no enhancement of the nonequilibrium population for state *d* and no corresponding change in stop-transfer efficiency. Previous experimental studies of stop-transfer efficiency involving relatively short CTL find no dependence of stop-transfer efficiency on translation rate (Hessa et al., 2003), as is consistent with the results in Figure 5D; experimental results for longer CTL that test the predicted kinetic effect upon slowing ribosomal translation would be of significant interest.

DISCUSSION

We have introduced a CG model for the direct simulation of cotranslational protein translocation and membrane integration on biological timescales. The model, which is based on MD simulations and limited experimental data, captures a striking array of experimentally observed features of integral membrane protein topogenesis and stop-transfer efficiency. The success of the model suggests that regulation of Sec-facilitated protein

translocation and membrane integration arises from simple features of the translocon machinery, including the confined geometry of the ribosome and translocon channel, conformational flexibility the translocon LG, and electrostatic and hydrophobic driving forces. Analysis of over 40,000 minute-timescale CG trajectories provides detailed insight into the mechanistic origin of the observed trends in protein targeting. In simulations of integral membrane protein topogenesis, the ensemble of CG trajectories suggests that the experimentally observed dependence of signal orientation on the ribosomal translation rate (Goder and Spiess, 2003) arises from the slow reorientation (i.e., flipping) of the SP in the confined environment of the translocon channel. In simulations of TM partitioning, the ensemble of CG trajectories suggests that the experimentally observed sigmoidal relationship between stop-transfer efficiency and the H-domain hydrophobicity (Hessa et al., 2005) arises from rapid local equilibration of the H-domain across the translocon LG. Finally, we utilize the CG model to predict the dependence of cotranslational protein stop-transfer efficiency on the ribosomal translation rate, protein nascent chain sequence, and protein CTL. The theoretical framework put forward in this paper provides a basis for testing and refining the mechanistic understanding of Sec-facilitated protein targeting.

EXPERIMENTAL PROCEDURES

Here, we present the CG model for direct simulation of cotranslational protein translocation and membrane integration. The model introduces necessary simplifications to reach the long timescales associated with these biological processes. It is parameterized using the results of MD simulations and transferable experimental data. Numerical testing, reported in Results and in the Extended Results, indicates that the CG model is consistent with independent experimental measurements of protein translocation and membrane integration and that reported conclusions are robust with respect to the details of the model parameterization.

The most aggressive simplification employed in the CG model is projection of the nascent protein dynamics onto the plane that passes along the translocon channel axis and between the helices of the LG (see Figure 1, as well the more detailed description below). The model includes explicit opening and closing of the translocon LG, which corresponds to the LG helices passing into and out of the plane of the nascent protein dynamics, but the nascent protein is itself confined to the planar subspace. This dimensionality reduction is necessary to make tractable the minute-timescale trajectories for protein translocation and membrane integration. Similar approaches are well established for the study of biomolecule transport and translocation systems. Planar models have been utilized for the theoretical analysis (Muthukumar, 1999; Panja et al., 2007; Sung and Park, 1996) and computer simulation (Chuang et al., 2002; Huopaniemi et al., 2006; Luo et al., 2007, 2008; Wei et al., 2007) of protein and DNA translocation through nanometer-lengthscale pores, and they have been used to investigate both thermodynamic and kinetic features of protein folding pathways (Dill et al., 1995; Go and Taketomi, 1978; Li and Cieplak, 1999). Even more simplified one-dimensional models of protein translocation have proved useful (Chauwin et al., 1998; Elston, 2000; Liebermeister et al., 2001; Simon et al., 1992). The success of such models follows from the pseudo-one-dimensional nature of pore-transport phenomena; kinetic bottlenecks are largely governed by progress transverse to the narrow pore, enabling dramatic simplification of other degrees of freedom. Although the CG model presented here is novel in that it explicitly describes translocon LG motions and ribosomal translation, it is based on the foundation of these earlier physical models.

Parameterization of the CG model utilizes MD simulations and transferable experimental data. Free energy calculations and direct MD simulations determine the energetics and timescales of LG opening, including the dependence of the LG energetics on the nascent-protein amino acid sequence; microsecond-timescale all-atom simulations and experimental measurements determine the diffusive timescale for the CG representation of the nascent protein; and experimental amino acid water/membrane transfer free energies determine the solvation energetics of the CG nascent protein residues.

Following initial parameterization, the CG model is left unchanged throughout the remainder of the study. Numerical tests indicate that the reported conclusions are robust with respect to geometric features of the translocon (Figure S8A) and the ribosome (Figure S8B), the timescales for translocon LG motion and nascent protein diffusion (Figures S9 and S10), features of the nascent protein sequence (Figures S2A, S2C, S2E, S2F, S5A, and S5B), and the effects of lumenal biasing factors, such as BiP (Figures S2B and S5E). These validation studies, as well as comparison of the simulations with experimental results (Figures 3, 5, S2, S3, and S5), suggest that the model captures the essential features of translocon-guided protein translocation and membrane integration.

Nonetheless, limitations of the CG model are emphasized from the outset. In addition to enforcing planar constraints on the motion of the nascent protein, the model provides a coarsened representation for nascent-protein, translocon, and membrane bilayer that includes only simple aspects of electrostatic and hydrophobic driving forces; potentially important details of residue-specific interactions are thus neglected (Dowhan and Bogdanov, 2009). Backbone interactions along the nascent protein chain are also neglected, such that effects due to the onset of nascent protein secondary structure are ignored, and effects due to translocon conformational changes other than LG motion are not explicitly included. Moreover, the possible roles of membrane-bound chaperones or oligomerization of the translocon channel (Hizlan et al., 2012) are not considered here. In principle, the CG model can be modified to incorporate greater accuracy and detail, as well as additional complexity and computational expense. In its current form, which is described in detail below, the model provides a minimalist description of Sec-facilitated protein translocation and membrane integration.

The System

The model employs CG particles, or beads, to describe the Sec translocon, protein nascent chain, hydrophobic membrane interior, and confinement effects due to the translating ribosome. The beads are constrained to the plane that lies normal to the lipid bilayer membrane and that bisects the translocon channel interior and the LG helices (Figure 1). CG beads corresponding to the residues of the translating nascent chain (Figure 1, inset) evolve subject to overdamped Brownian dynamics, whereas beads representing the Sec translocon (light and dark green) and the docked ribosome (brown) are fixed with respect to the membrane bilayer. To explicitly incorporate the conformational gating of the translocon LG helices, beads representing the LG helices (dark green) undergo stochastic transitions between closed-state interactions, which occlude the passage of the nascent chain from the Sec channel to the membrane interior, and open-state interactions, for which the steric barrier to membrane integration is removed. Structural features of the channel and ribosomal confinement are obtained from crystallographic and electron microscopy studies (Beckmann et al., 2001; Van den Berg et al., 2004). The positions for the translocon and ribosome beads are reported in Table S1.

Interactions

We employ a CG bead diameter of $\sigma = 8$ Å, which is typical of the Kuhn length for polypeptide chains (Hanke et al., 2010; Staple et al., 2008); the protein nascent protein chain is thus modeled as a freely jointed chain with each CG bead corresponding to approximately three amino acid residues. Bonding interactions between neighboring beads in the nascent chain are described using the finite extension nonlinear elastic (FENE) potential (Kremer and Grest, 1990), $U(r) = -1/2\, kR_0^2 \ln(1 - r^2/R_0^2)$, where $k = 7\varepsilon/\sigma^2$, $R_0 = 2\sigma$, and $\varepsilon = 0.833 k_B T$; all simulations are performed using $T = 300$ K. The bonding interactions are sufficiently strong to avoid self-crossing of the protein nascent chain.

For the description of nonbonded interactions, the CG beads are categorized into various types. For the protein nascent chain, the CG bead types correspond to positively charged (R), negatively charged (E), neutral-hydrophobic (L), neutral-hydrophilic (Q), mildly hydrophobic (V), amphiphilic (P),

and variable-hydrophobic (X) groups of amino acid residues. Additional CG bead types correspond to residues of the ribosome, residues for the translocon LG in the closed state (LG_c), residues for the translocon LG in the open state (LG_o), and residues for the translocon that are not part of the LG (LG_n).

Short-ranged nonbonding interactions are modeled using the Lennard-Jones (LJ) potential energy function,

$$U_{LJ}(r) = \begin{cases} 4\varepsilon_{lj}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \varepsilon_{cr}, & r_{cl} < r \leq r_{cr} \\ 0, & \text{otherwise} \end{cases}, \quad \text{(Equation 2)}$$

where the constant ε_{cr} ensures that the pairwise interaction vanishes at r_{cr}. For each pair of CG bead types, the corresponding LJ parameters are reported in Table S2. For the nonbonding interactions among the beads of the protein nascent chain and between beads of the nascent chain and the ribosome, the LJ parameters correspond to soft-walled, excluded volume interactions (Weeks et al., 1971). Weak attractive interactions account for the affinity of the protein nascent chain for the LG helices of the translocon, as has been observed in crosslinking experiments (Plath et al., 1998). For the open state of the LG, repulsions between the LG and protein nascent chain beads are truncated to allow the peptide to laterally exit the translocon channel.

Pairwise Coulombic interactions are modeled using the Debye-Hückel potential, $U_{DH}(r) = \sigma q_1 q_2 (\beta r)^{-1} \exp[-r/\kappa]$, where q_1 and q_2 are the charges for the various CG beads (Table S3). We employ a Debye length of $\kappa = 1.4\sigma$ that is typical for electrostatic screening under physiological conditions. Two additional charges are included to model charge distribution among the residues of the translocon; a charge of $q = -2$ and $q = 2$ are included on the first and fourth beads of the LG, where the LG beads are ordered with respect to their distance from the cytosol. The justification for the LG bead charges is discussed in Extended Experimental Procedures. When the LG is in the open state, the electrostatic potential between the beads on the LG and on the protein nascent chain is capped from below to avoid the singularity in the Debye-Hückel potential, such that

$$U(r) = \begin{cases} U_{DH}(r), & r > \sigma \\ U_{DH}(\sigma), & \text{otherwise.} \end{cases} \quad \text{(Equation 3)}$$

Solvation energetics for each CG bead are described using the position-dependent potential energy function

$$U_{solv}(x,y) = gS(x;\phi_x,\psi_x)\left[1 - S(y;\phi_y,\psi_y)\right], \quad \text{(Equation 4)}$$

where x and y are the Cartesian coordinates for the CG bead (Figure S7), and g is the corresponding water-membrane transfer FE (Table S3). Smooth transitions for the bead solvation energy upon moving from aqueous to membrane environments are achieved using the switching function

$$S(x;\phi,\psi) = \frac{1}{4}\left(1 + \tanh\frac{x-\phi}{b}\right)\left(1 - \tanh\frac{x-\psi}{b}\right), \quad \text{(Equation 5)}$$

where the switching lengthscale is $b = 0.25\sigma$. The parameters that describe the switching between the aqueous and membrane regions of the system are $\phi_x = -2.0\sigma$, $\psi_x = 2.0\sigma$, $\phi_y = -1.5\sigma$, $\psi_y = 1.5\sigma$.

Dynamics

The time-evolution of the system is modeled using a combination of Brownian dynamics for the nascent protein chain and stochastic opening and closing of the translocon LG. The off-lattice nascent chain dynamics is evolved using the first-order Euler integrator (Stoer and Bulirsch, 2002)

$$x_i(t+\Delta t) = x_i(t) - \beta D\frac{\partial V(\mathbf{x}(t))}{\partial x_i}\Delta t + \sqrt{2D\Delta t}\eta_i, \quad \text{(Equation 6)}$$

where $x_i(t)$ is a Cartesian degree of freedom for the nascent chain at time t, $V(\mathbf{x}(t))$ is the potential energy function for the full system, D is the isotropic diffusion constant for the CG beads, $\beta = (k_B T)^{-1}$, and η is a random number drawn from the Gaussian distribution with zero mean and unit variance. As is described in Extended Experimental Procedures, a CG bead diffusion constant of $D = 758.7\ \text{nm}^2/\text{s}$ reproduces experimentally observed timescales for nascent chain diffusion through the translocon channel (Elston, 2000; Matlack et al., 1999) and is consistent with microsecond all-atom MD simulations. With this diffusion constant and the previously described interaction parameters, Equation 6 can be stably integrated with a timestep of $\Delta t = 100$ ns.

At every simulation timestep, the probability of LG opening/closing is $p_{open/close} = k_{open/close}\Delta t$, where

$$k_{open} = \frac{1}{\tau_{LG}}\frac{\exp(-\beta\Delta G_{tot})}{1+\exp(-\beta\Delta G_{tot})}, \quad \text{(Equation 7)}$$

and

$$k_{close} = \frac{1}{\tau_{LG}}\frac{1}{1+\exp(-\beta\Delta G_{tot})}. \quad \text{(Equation 8)}$$

Here, τ_{LG} corresponds to the timescale for attempting LG opening or closing events, and ΔG_{tot} is the FE cost associated with LG opening. As is described in Extended Experimental Procedures, the calculation of ΔG_{tot}, as well as the dependence of this FE cost on the nascent chain contents of the translocon channel, is based on MD simulations of the channel/peptide-substrate/membrane system (Zhang and Miller, 2010). The timescale $\tau_{LG} = 500$ ns is likewise determined from MD simulations (Zhang and Miller, 2010). Equations 6, 7, and 8 satisfy detailed balance, ensuring that the CG dynamics is consistent with equilibrium Boltzmann statistics.

Modeling Translation

Ribosomal translation is directly modeled in the CG simulations via growth of the nascent chain at the ribosome exit channel (Figure 1, inset, red). The C terminus of the protein nascent chain is held fixed at the exit channel throughout translation, and beads are sequentially added at the C-terminal tail, elongating the protein nascent chain. Upon completion of translation, the nascent chain is released from the exit channel, and the small subunit of the ribosome dissociates from the cytosolic mouth of translocon (Heritage and Wonderlin, 2001; Seiser and Nicchitta, 2000); we model ribosomal dissociation by eliminating interactions associated with the ribosome CG beads. Ribosomal translation proceeds at a pace of approximately 10–20 amino acid residues per second (res/s) (Bilgin et al., 1992; Boehlke and Friesen, 1975), although this rate can be reduced approximately 4-fold upon addition of cycloheximide (Abou Elela and Nazar, 1997; Goder and Spiess, 2003); we thus consider ribosomal translation rates in the range of 6–24 res/s (2–8 beads/s) in the current study.

The binding immunoglobulin protein (BiP) is an essential component of the eukaryotic Sec translocon machinery (Brodsky et al., 1995). In Explicit Modeling of Lumenal BiP, we consider the explicit inclusion of BiP binding within the CG model and show that it gives rise to only modest effects in the calculated results for protein translation and membrane integration. Unless otherwise stated, explicit BiP binding is not included in the reported simulation results.

SUPPLEMENTAL INFORMATION

Supplemental Information includes ten figures, three tables, and five movies and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2012.08.039.

LICENSING INFORMATION

ACKNOWLEDGMENTS

This research was supported in part by the U.S. Office of Naval Research (USONR) under Grant No. N00014-10-1-0884, and T.F.M. acknowledges an Alfred P. Sloan Foundation fellowship. Computational resources were

provided by the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, and by the National Science Foundation under Grant No. CHE-1040558. We additionally acknowledge use of the Anton super-computer system that is hosted by the National Resource for Biomedical Supercomputing (NRBSC) at the Pittsburgh Supercomputing Center (PSC), with funding from the National Institute of General Medical Sciences under grant RC2GM093307.

Received: March 30, 2012
Revised: July 21, 2012
Accepted: August 31, 2012
Published online: October 18, 2012

REFERENCES

Abou Elela, S., and Nazar, R.N. (1997). Role of the 5.8S rRNA in ribosome translocation. Nucleic Acids Res. *25*, 1788–1794.

Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. (2001). Architecture of the protein-conducting channel associated with the translating 80S ribosome. Cell *107*, 361–372.

Beltzer, J.P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H.P., and Spiess, M. (1991). Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence. J. Biol. Chem. *266*, 973–978.

Bieker, K.L., and Silhavy, T.J. (1990). PrlA (SecY) and PrlG (SecE) interact directly and function sequentially during protein translocation in E. coli. Cell *61*, 833–842.

Bilgin, N., Claesens, F., Pahverk, H., and Ehrenberg, M. (1992). Kinetic properties of Escherichia coli ribosomes with altered forms of S12. J. Mol. Biol. *224*, 1011–1027.

Boehlke, K.W., and Friesen, J.D. (1975). Cellular content of ribonucleic acid and protein in Saccharomyces cerevisiae as a function of exponential growth rate: calculation of the apparent peptide chain elongation rate. J. Bacteriol. *121*, 429–433.

Bonardi, F., Halza, E., Walko, M., Du Plessis, F., Nouwen, N., Feringa, B.L., and Driessen, A.J. (2011). Probing the SecYEG translocation pore size with preproteins conjugated with sizable rigid spherical molecules. Proc. Natl. Acad. Sci. USA *108*, 7775–7780.

Bondar, A.N., del Val, C., Freites, J.A., Tobias, D.J., and White, S.H. (2010). Dynamics of SecY translocons with translocation-defective mutations. Structure *18*, 847–857.

Brodsky, J.L., Goeckeler, J., and Schekman, R. (1995). BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. Proc. Natl. Acad. Sci. USA *92*, 9643–9646.

Chauwin, J.F., Oster, G., and Glick, B.S. (1998). Strong precursor-pore interactions constrain models for mitochondrial protein import. Biophys. J. *74*, 1732–1743.

Cheng, Z., and Gilmore, R. (2006). Slow translocon gating causes cytosolic exposure of transmembrane and lumenal domains during membrane protein integration. Nat. Struct. Mol. Biol. *13*, 930–936.

Chuang, J., Kantor, Y., and Kardar, M. (2002). Anomalous dynamics of translocation. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *65*, 011802.

Crowley, K.S., Reinhart, G.D., and Johnson, A.E. (1993). The signal sequence moves through a ribosomal tunnel into a noncytoplasmic aqueous environment at the ER membrane early in translocation. Cell *73*, 1101–1115.

Crowley, K.S., Liao, S., Worrell, V.E., Reinhart, G.D., and Johnson, A.E. (1994). Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore. Cell *78*, 461–471.

Denzer, A.J., Nabholz, C.E., and Spiess, M. (1995). Transmembrane orientation of signal-anchor proteins is affected by the folding state but not the size of the N-terminal domain. EMBO J. *14*, 6311–6317.

Devaraneni, P.K., Conti, B., Matsumura, Y., Yang, Z., Johnson, A.E., and Skach, W.R. (2011). Stepwise insertion and inversion of a type II signal anchor sequence in the ribosome-Sec61 translocon complex. Cell *146*, 134–147.

Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D., and Chan, H.S. (1995). Principles of protein folding—a perspective from simple exact models. Protein Sci. *4*, 561–602.

Do, H., Falcone, D., Lin, J., Andrews, D.W., and Johnson, A.E. (1996). The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. Cell *85*, 369–378.

Dowhan, W., and Bogdanov, M. (2009). Lipid-dependent membrane protein topogenesis. Annu. Rev. Biochem. *78*, 515–540.

Duong, F., and Wickner, W. (1998). Sec-dependent membrane protein biogenesis: SecYEG, preprotein hydrophobicity and translocation kinetics control the stop-transfer function. EMBO J. *17*, 696–705.

Egea, P.F., and Stroud, R.M. (2011). Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. Proc. Natl. Acad. Sci. USA *107*, 17182–17187.

Elston, T.C. (2000). Models of post-translational protein translocation. Biophys. J. *79*, 2235–2251.

Frauenfeld, J., Gumbart, J., Sluis, E.O., Funes, S., Gartmann, M., Beatrix, B., Mielke, T., Berninghausen, O., Becker, T., Schulten, K., and Beckmann, R. (2011). Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. Nat. Struct. Mol. Biol. *18*, 614–621.

Garrison, J.L., Kunkel, E.J., Hegde, R.S., and Taunton, J. (2005). A substrate-specific inhibitor of protein translocation into the endoplasmic reticulum. Nature *436*, 285–289.

Go, N., and Taketomi, H. (1978). Respective roles of short- and long-range interactions in protein folding. Proc. Natl. Acad. Sci. USA *75*, 559–563.

Goder, V., and Spiess, M. (2001). Topogenesis of membrane proteins: determinants and dynamics. FEBS Lett. *504*, 87–93.

Goder, V., and Spiess, M. (2003). Molecular mechanism of signal sequence orientation in the endoplasmic reticulum. EMBO J. *22*, 3645–3653.

Gumbart, J., and Schulten, K. (2006). Molecular dynamics studies of the archaeal translocon. Biophys. J. *90*, 2356–2367.

Gumbart, J., and Schulten, K. (2007). Structural determinants of lateral gate opening in the protein translocon. Biochemistry *46*, 11147–11157.

Gumbart, J., Chipot, C., and Schulten, K. (2011). Free-energy cost for translocon-assisted insertion of membrane proteins. Proc. Natl. Acad. Sci. USA *108*, 3596–3601.

Haider, S., Hall, B.A., and Sansom, M.S. (2006). Simulations of a protein translocation pore: SecY. Biochemistry *45*, 13018–13024.

Hanke, F., Serr, A., Kreuzer, H.J., and Netz, R.R. (2010). Stretching single polypeptides: the effect of rotational constraints in the backbone. Europhys. Lett. *92*, 53001.

Harley, C.A., Holt, J.A., Turner, R., and Tipper, D.J. (1998). Transmembrane protein insertion orientation in yeast depends on the charge difference across transmembrane segments, their total hydrophobicity, and its distribution. J. Biol. Chem. *273*, 24963–24971.

Heinrich, S.U., Mothes, W., Brunner, J., and Rapoport, T.A. (2000). The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. Cell *102*, 233–244.

Heritage, D., and Wonderlin, W.F. (2001). Translocon pores in the endoplasmic reticulum are permeable to a neutral, polar molecule. J. Biol. Chem. *276*, 22655–22662.

Hessa, T., Monné, M., and von Heijne, G. (2003). Stop-transfer efficiency of marginally hydrophobic segments depends on the length of the carboxy-terminal tail. EMBO Rep. *4*, 178–183.

Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S.H., and von Heijne, G. (2005). Recognition of transmembrane helices by the endoplasmic reticulum translocon. Nature *433*, 377–381.

Hessa, T., Meindl-Beinker, N.M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S.H., and von Heijne, G. (2007). Molecular code for transmembrane-helix recognition by the Sec61 translocon. Nature *450*, 1026–1030.

Higy, M., Gander, S., and Spiess, M. (2005). Probing the environment of signal-anchor sequences during topogenesis in the endoplasmic reticulum. Biochemistry *44*, 2039–2047.

Hikita, C., and Mizushima, S. (1992). Effects of total hydrophobicity and length of the hydrophobic domain of a signal peptide on in vitro translocation efficiency. J. Biol. Chem. *267*, 4882–4888.

Hizlan, D., Robson, A., Whitehouse, S., Gold, V.A., Vonck, J., Mills, D., Kühlbrandt, W., and Collinson, I. (2012). Structure of the SecY complex unlocked by a preprotein mimic. Cell Rep. *1*, 21–28.

Huopaniemi, I., Luo, K., Ala-Nissila, T., and Ying, S.-C. (2006). Langevin dynamics simulations of polymer translocation through nanopores. J. Chem. Phys. *125*, 124901.

Jungnickel, B., and Rapoport, T.A. (1995). A posttargeting signal sequence recognition event in the endoplasmic reticulum membrane. Cell *82*, 261–270.

Junne, T., Schwede, T., Goder, V., and Spiess, M. (2007). Mutations in the Sec61p channel affecting signal sequence recognition and membrane protein topology. J. Biol. Chem. *282*, 33201–33209.

Kim, S.J., Mitra, D., Salerno, J.R., and Hegde, R.S. (2002). Signal sequences control gating of the protein translocation channel in a substrate-specific manner. Dev. Cell *2*, 207–217.

Junne, T., Kocik, L., and Spiess, M. (2010). The hydrophobic core of the Sec61 translocon defines the hydrophobicity threshold for membrane integration. Mol. Biol. Cell *21*, 1662–1670.

Kremer, K., and Grest, G.S. (1990). Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. J. Chem. Phys. *92*, 5057–5086.

Li, M.S., and Cieplak, M. (1999). Folding in two-dimensional off-lattice models of proteins. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics *59*, 970–976.

Liebermeister, W., Rapoport, T.A., and Heinrich, R. (2001). Ratcheting in post-translational protein translocation: a mathematical model. J. Mol. Biol. *305*, 643–656.

Luo, K., Ala-Nissila, T., Ying, S.-C., and Bhattacharya, A. (2007). Influence of polymer-pore interactions on translocation. Phys. Rev. Lett. *99*, 148102.

Luo, K., Ala-Nissila, T., Ying, S.-C., and Bhattacharya, A. (2008). Sequence dependence of DNA translocation through a nanopore. Phys. Rev. Lett. *100*, 058101.

Maifeld, S.V., MacKinnon, A.L., Garrison, J.L., Sharma, A., Kunkel, E.J., Hegde, R.S., and Taunton, J. (2011). Secretory protein profiling reveals TNF-α inactivation by selective and promiscuous Sec61 modulators. Chem. Biol. *18*, 1082–1088.

Matlack, K.E., Misselwitz, B., Plath, K., and Rapoport, T.A. (1999). BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane. Cell *97*, 553–564.

Meindl-Beinker, N.M., Lundin, C., Nilsson, I., White, S.H., and von Heijne, G. (2006). Asn- and Asp-mediated interactions between transmembrane helices during translocon-mediated membrane protein assembly. EMBO Rep. *7*, 1111–1116.

Muthukumar, M. (1999). Polymer translocation through a hole. J. Chem. Physiol. *111*, 10371–10374.

Ojemalm, K., Higuchi, T., Jiang, Y., Langel, U., Nilsson, I., White, S.H., Suga, H., and von Heijne, G. (2011). Apolar surface area determines the efficiency of translocon-mediated membrane- protein integration into the endoplasmic reticulum. Proc. Natl. Acad. Sci. USA *108*, E359–E364.

Panja, D., Barkema, G.T., and Ball, R.C. (2007). Anomalous dynamics of unbiased polymer translocation through a narrow pore. J. Phys. Condens. Matter *19*, 432202–432202.

Park, E., and Rapoport, T.A. (2011). Preserving the membrane barrier for small molecules during bacterial protein translocation. Nature *473*, 239–242.

Parks, G.D., and Lamb, R.A. (1991). Topology of eukaryotic type II membrane proteins: importance of N-terminal positively charged residues flanking the hydrophobic domain. Cell *64*, 777–787.

Plath, K., Mothes, W., Wilkinson, B.M., Stirling, C.J., and Rapoport, T.A. (1998). Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. Cell *94*, 795–807.

Rapoport, T.A., Goder, V., Heinrich, S.U., and Matlack, K.E. (2004). Membrane-protein integration and the role of the translocation channel. Trends Cell Biol. *14*, 568–575.

Rutkowski, D.T., Lingappa, V.R., and Hegde, R.S. (2001). Substrate-specific regulation of the ribosome- translocon junction by N-terminal signal sequences. Proc. Natl. Acad. Sci. USA *98*, 7823–7828.

Rychkova, A., Vicatos, S., and Warshel, A. (2010). On the energetics of translocon-assisted insertion of charged transmembrane helices into membranes. Proc. Natl. Acad. Sci. USA *107*, 17598–17603.

Seiser, R.M., and Nicchitta, C.V. (2000). The fate of membrane-bound ribosomes following the termination of protein synthesis. J. Biol. Chem. *275*, 33820–33827.

Shaw, A.S., Rottier, P.J., and Rose, J.K. (1988). Evidence for the loop model of signal-sequence insertion into the endoplasmic reticulum. Proc. Natl. Acad. Sci. USA *85*, 7592–7596.

Simon, S.M., Peskin, C.S., and Oster, G.F. (1992). What drives the translocation of proteins? Proc. Natl. Acad. Sci. USA *89*, 3770–3774.

Smith, M.A., Clemons, W.M., Jr., DeMars, C.J., and Flower, A.M. (2005). Modeling the effects of prl mutations on the Escherichia coli SecY complex. J. Bacteriol. *187*, 6454–6465.

Sokal, R.R., and Rohlf, F.J. (1994). Biometry: The Principles and Practices of Statistics in Biological Research, Third Edition (New York: W.H. Freeman).

Staple, D.B., Payne, S.H., Reddin, A.L.C., and Kreuzer, H.J. (2008). Model for stretching and unfolding the giant multidomain muscle protein using single-molecule force spectroscopy. Phys. Rev. Lett. *101*, 248301.

Stoer, J., and Bulirsch, R. (2002). Introduction to Numerical Analysis (New York: Springer).

Sung, W., and Park, P.J. (1996). Polymer translocation through a pore in a membrane. Phys. Rev. Lett. *77*, 783–786.

Tian, P., and Andricioaei, I. (2006). Size, motion, and function of the SecY translocon revealed by molecular dynamics simulations with virtual probes. Biophys. J. *90*, 2718–2730.

Tsukazaki, T., Mori, H., Fukai, S., Ishitani, R., Mori, T., Dohmae, N., Perederina, A., Sugita, Y., Vassylyev, D.G., Ito, K., and Nureki, O. (2008). Conformational transition of Sec machinery inferred from bacterial SecYE structures. Nature *455*, 988–991.

Van den Berg, B., Clemons, W.M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S.C., and Rapoport, T.A. (2004). X-ray structure of a protein-conducting channel. Nature *427*, 36–44.

von Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. EMBO J. *5*, 3021–3027.

Wahlberg, J.M., and Spiess, M. (1997). Multiple determinants direct the orientation of signal-anchor proteins: the topogenic role of the hydrophobic signal domain. J. Cell Biol. *137*, 555–562.

Weeks, J.D., Chandler, D., and Andersen, H.C. (1971). Role of repulsive forces in determining the equilibrium structure of simple liquids. J. Chem. Phys. *54*, 5237–5247.

Wei, D., Yang, W., Jin, X., and Liao, Q. (2007). Unforced translocation of a polymer chain through a nanopore: the solvent effect. J. Chem. Phys. *126*, 204901.

Zhang, B., and Miller, T.F., 3rd. (2010). Hydrophobically stabilized open state for the lateral gate of the Sec translocon. Proc. Natl. Acad. Sci. USA *107*, 5399–5404.

Zhang, B., and Miller, T.F., 3rd. (2012). Direct simulation of early-stage sec-facilitated protein translocation. J. Am. Chem. Soc. *134*, 13700–13707.

Zimmer, J., Nam, Y., and Rapoport, T.A. (2008). Structure of a complex of the ATPase SecA and the protein-translocation channel. Nature *455*, 936–943.